QA

Measuring the Competence of Healthcare Providers

by Neeraj Kak, Bart Burkhalter, and Merri-Ann Cooper

Executive Summary

Competence encompasses knowledge, skills, abilities, and traits. It is gained in the healthcare professions through pre-service education, in-service training, and work experience. Competence is a major determinant of provider performance as represented by conformance with various clinical, non-clinical, and interpersonal standards. Measuring competence is essential for determining the ability and readiness of health workers to provide quality services. Although competence is a precursor to doing the job right, measuring performance periodically is also crucial to determine whether providers are using their competence on the job. A provider can have the knowledge and skill, but use it poorly because of individual factors (abilities, traits, goals, values, inertia, etc.) or external factors (unavailability of drugs, equipment, organizational support, etc.).

This paper provides a framework for understanding the key factors that affect provider competence. Different methods for measuring competence are discussed, as are criteria for selecting measurement methods. Also, evidence from various research studies on measuring the effectiveness of different assessment techniques is presented.

Introduction

Understanding the causes of poor performance of healthcare providers in both developed and developing countries is crucial to high quality healthcare. To the extent poor performance is caused by low competence, improving competency would improve performance. But how are performance and competence linked, and how well can we measure competence?

QA□□□
QUALITY ASSURANCE PROJECT

USAID

The Quality Assurance Project endeavors to improve healthcare provider performance. The QA Project developed this paper on measuring competence to guide healthcare systems in improving their performance through better hiring, job restructuring, re-organization, and the like. The paper focuses on competence and reviews several studies that have contributed to the understanding of competency in medical education and healthcare settings. Little research exists—and more is needed—on measuring and improving competency in developing country healthcare settings.

Limits of this paper

The conclusions about competence measurement are largely drawn from studies conducted in the developed world with healthcare students, nurses, physicians, and other healthcare workers. Very few studies have been designed and conducted in developing countries on measuring competence and the relationship between competence and provider behavior. However, the measurement issues involved in the assessment of competence, including

CONTENTS

validity and reliability, are relevant within any cultural context. The findings relevant to these issues (e.g., what types of measures are most reliable) likely apply in any country. However, the applicability of the other findings and recommendations should be evaluated for relevance to the specific situation. For example, there is extensive evidence from developed countries that detailed and immediate feedback on performance improves both learning and later performance. Will such feedback procedures be equally effective in all cultures? The reader is cautioned against assuming that all of the conclusions in this paper, coming from one cultural context, will necessarily apply to other contexts.

Second, the literature reviewed in this paper concerns the clinical competence of individuals. Research studies on the evaluation of team performance exist, but were not reviewed for this paper.

Last, self-assessment, an emerging method for measuring competence, is addressed in a forthcoming *Issue Paper:* "How Can Self-Assessment Improve the Quality of Healthcare?"

What is competence?

Competence refers to a person's underlying characteristics that are causally related to job performance (Boyatzis 1982). Competence is defined in the context of particular *knowledge, traits, skills,* and *abilities.* Knowledge involves understanding facts and procedures. Traits are personality characteristics (e.g., self-control, self-confidence) that predispose a person to behave or respond in a certain way. Skill is the capacity to perform specific actions: a person's skill is a function of both knowledge and the particular strategies used to apply knowledge. Abilities are the attributes that a person has inherited or acquired through previous experience and brings to a new task (Landy 1985): they are more fundamental and stable than knowledge and skills (Fleishman and Bartlett 1969).

Competence can be defined as the ability to perform a specific task in a manner that yields desirable outcomes. This definition implies the ability to apply knowledge, skills, and abilities successfully to new situations as well as to familiar tasks for which prescribed standards exist (Lane and Ross 1998). Health workers acquire competence over time (Benner 1984). Typically, pre-service education or an initial training opportunity creates a novice who, after additional training and hands-on experience, reaches a level that can be certified as competent. Although competence is considered to be a major milestone in professional development, it is not the final point. That comes with proficiency, and the ultimate status of expert comes after many years of experience and professional growth (Benner 1984).

Competence is one of many determinants of performance. The relationship between competence (can do) and performance (does do) is complex: the first does not always predict the second (Southgate and Dauphinee 1998). Obviously, less competent providers are less likely to provide quality services, and healthcare providers must have the competencies necessary to perform their jobs according to standards in order to provide quality services.

Attempts are sometimes made to measure competence in terms of performance. However, competence should not be inferred from performance (While 1994). While competence is defined in terms of someone's capacity to perform, performance is the resulting behavior. "Performance is something that people actually do and can be observed. By definition, it includes only those actions or behaviors that are relevant to the organization's goals and that can be scaled (measured) in terms of each person's proficiency (that is, level of contribution). Performance is what the organization hires one to do, and do well" (Campbell et al. 1993).

Abbreviations	
CE	Continuing education
CEF	Clinical evaluation form
CHR	Center for Health Research
CPR	Cardio-pulmonary resuscitation
CST	Clinical simulation testing
DSS	Decision support systems
IUD	Intrauterine device
IV	Intravenous needle
JCAHO	Joint Commission on the Accreditation of Healthcare Organizations
KTS	Knowledge test of skills
MSE	Multiple-station examination
OACA	Objective Assessment of Competence Achievement
OSCE	Objective structured clinical examination
PBL	Problem-based learning
PBT	Performance-based test
PVA	Practice video assessment
SAQ	Self-assessment questionnaire
UAE	United Arab Emirates
UK	United Kingdom
USAID	U.S. Agency for International Development

Measuring competence

Why measure competence?

There are many good reasons for measuring competence. Ministries of health, professional organizations, and healthcare organizations must ensure that appropriate expectations for competence are set and that their staff perform to standard. Healthcare organizations must meet certain criteria to provide services. These organizations—through certification, licensure, and/or accreditation—are able to exert control on health providers and, as a result, to influence the quality of care. Although most health providers must demonstrate minimum competence during training to move up to the next level or graduate from a course, not all healthcare organizations assess job- or skill-specific competencies before offering employment. Reasons why healthcare organizations should measure competence include:

Healthcare reform: The increasing complexities of healthcare delivery and changing market conditions have forced health policy-makers to promote the assessment of initial competence of students and new graduates and the continuing competence of experienced and certified practitioners (Lenburg 1999). In the United States, this has led to various reforms affecting teaching and the methods used to assess students' actual competence. The Joint Commission on the Accreditation of Healthcare Organizations (JCAHO) has required specific validation of the competence of healthcare providers for institutional accreditation (JCAHO 1996). With rapid technological and scientific innovations, U.S. employers are spending substantial resources to assure competencies of new and existing health staff. These health institutions are using specific standardized tests to document the readiness of employees for full practice responsibilities (Lenburg 1999).

Organizational performance: Are providers effectively treating their clients? And if not, why? Is the problem a lack of competence? Healthcare organizations need to assess individual and organizational performance periodically to assess the efficacy of their services. The results help healthcare organizations determine whether they need to design training and/or continuing education interventions for improving provider performance.

Comparing assessments of competence and job performance may indicate the extent to which the organization provides the support needed for quality care. High competency and low performance may signal that an organization is not providing the needed resources, has not clarified standards of care, or is not rewarding effective performance or correcting poor performance.

Liability and ethics: Healthcare organizations are responsible for the quality of care their staff provide and consequently must ensure that their staffs are competent and can meet standards for the provision of care. Assessing providers' competence periodically enables healthcare organizations to meet this crucial responsibility.

Risk management: Competency assessments can be used to monitor organization-wide knowledge of policies and procedures related to high-risk areas. Feedback from these assessments can be used for training and continuing education of providers and to improve overall organizational performance.

Certification and recertification of providers: Competency assessment is an integral part of the certification and recertification processes of service providers. For example, recertification programs in the U.S. use examinations and performance assessments as "snapshots" of competence every seven to ten years (Bashook and Parboosingh 1998).

Documenting competence is becoming essential—not optional—and is likely to become mandatory in the near future for initial and continuing licensure and certification, and perhaps even for employment (Lenburg 1999).

Planning for new services: Competency assessment can help managers identify providers who are competent to provide a new clinical service, providers who need improvements in specific knowledge or skill areas when a new service is offered, and providers who are ready to act as mentors of newly trained providers.

Measuring training outcomes: Competency assessment can determine the efficacy of training interventions in closing knowledge and skill gaps and to assess and improve training. Low scores on competence assessments after training may indicate that the training was ineffective, poorly designed, poorly presented, or inappropriate. Trainers can use this information to improve training content or delivery. If the assessments aim to improve specific components of training, the trainer may be able to determine where more information is needed, which exercises require clarification, or if more time is required to cover a topic (Smith and Merchant 1990).

Selection of new staff: Competency assessment is useful when recruiting new staff to ensure they can do the job they are hired to do or could do it with reasonable orientation/training.

Individual performance improvement: Competency assessment can play an important role in an organization's performance improvement initiatives. Assessment results can identify gaps in knowledge and skills, and guide managers in setting appropriate training or other remedial interventions targeting individual providers or groups of providers.

Supervision: Competency assessments can guide healthcare managers in providing performance improvement feedback to healthcare providers.

Restrictions on competency assessments

For budgetary or other reasons, many health organizations may not routinely measure competence across the breadth of workplace tasks until a performance problem becomes apparent. Also, some competencies may be difficult to describe precisely or evaluate accurately. In such cases, it is essential that health programs link competencies with specific performance indicators as a proxy for measuring competence (Lane and Ross 1998).

Which competencies should be measured?

Health workers need a large number of competencies for providing quality services. Benner (1984) and Fenton (1985) have proposed several domains with specific competencies that are critical for nursing care (see Table 1). Although these domains have been defined from a nursing practice perspective, they apply equally to other types of health workers. Some of these competencies affect the quality of care directly and others indirectly. While health workers can gain knowledge about various competencies during pre-service education, skills related to these competencies are further advanced during practicum or on the job. In addition, mentors and preceptors can further assist to improve health worker competency.

Conceptual framework

Competency is defined as the ability of a health worker to perform according to predefined standards. Competency is developed through pre-service education, in-service training, hands-on experience, and the assistance of mentors and preceptors.

Competency measurement is critical to ensuring that all employees are competent to perform their assigned duties

---

Model ■ Novice to Expert

Patricia Benner (1984) provides a framework describing knowledge embedded in nursing practice that accrues over time. The model differentiates theoretical knowledge from practical knowledge acquired from clinical practice. The five levels of skill acquisition are:

Novice: The novice has no background or experience in his or her area. Rules and objective attributes are applied without an understanding or knowledge of the context of the situation. Nursing students are novices in nursing; one becomes a novice whenever he or she is placed in an unfamiliar area of practice.

Advanced Beginner: The advanced beginner demonstrates marginally acceptable performance based on experience acquired under the mentoring of a more experienced nurse or a teacher. The practice of nursing is rules-based and oriented toward completion of tasks. The larger context of the situation is difficult to grasp at this stage. There is a concern for good management of skills and time, but the need for guidance and assistance remains.

Competent: Competent nurses are able to differentiate between the aspects of the current situation and those of the future and can select those aspects that are important. The focus on good management of time skills remains, but the sense of responsibility is higher. However, they may have an unrealistic concept of what they can actually handle.

Proficient: Proficient nurses are able to see the whole situation in context and can apply knowledge to clinical practice, identifying the most salient aspects and differentiating them from those that are less important. Actions are intuitive and skilled. They have confidence in their own knowledge and abilities, focus less on rules and time management.

Expert: The expert nurse is able to focus intuitively on solutions to situations without having to explore alternatives. This ability is based on a rich experiential background. Focus is on meeting patient needs and concerns to the point of being an advocate for the patient and care. The focus on self and one's own performance is diminished.

---

and responsibilities, and to meet performance standards. Competence of all staff should be assessed prior to employment, during the orientation period, and at least annually thereafter. A well-defined job description and use of appropriate competency assessment tools go a long way in ensuring that competent providers are recruited to begin with and that through periodic assessments appropriate remedial actions are taken to close any competence gaps.

Periodic competence assessments should be considered for those areas that are considered *low-volume, high-risk,* or *critical* (Centra Health 1999). Low-volume competencies are those that occur so infrequently that they need to be assessed at least annually to ensure that providers are still able to perform these duties. High-risk competencies are

Table 1 ■ Domains of Nursing Care and Key Competencies

Domain	Key Competencies
The helping role	■ The healing relationship ■ Providing comfort measures and preserving personhood in the face of pain and/or extreme breakdown ■ Maximizing the patient's participation and control in her/his own recovery ■ Providing informational and emotional support to patient's family ■ Guiding patient through emotional and developmental change; providing new options, closing old ones
The diagnostic and monitoring function	■ Detecting and documenting significant changes in patient's condition ■ Providing an early warning signal: anticipating breakdown and deterioration prior to explicit, confirming diagnostic signs ■ Anticipating problems ■ Understanding the particular demands and experiences of an illness: anticipating patient care needs ■ Assessing the patient's potential for wellness and for responding to various treatment strategies
Administering and monitoring therapeutic interventions and regimens	■ Starting and maintaining intravenous therapy with minimal risks and complications ■ Administering medications accurately and safely: monitoring untoward effects, reactions, therapeutic responses, toxicity, incompatibilities ■ Creating a wound management strategy that fosters healing, comfort, and appropriate drainage
Effective management of rapidly changing situations	■ Skilled performance in extreme life-threatening emergencies: rapid grasp of a problem ■ Contingency management: rapid matching of demands and resources in emergency situations ■ Identifying and managing patient crisis until physician assistance is available
The teaching-coaching function	■ Timing: capturing a patient's readiness to learn ■ Assisting patients in integrating the implications of illness and recovery into their lifestyles ■ Eliciting and understanding the patient's interpretation of his or her illness ■ Providing an interpretation of the patient's condition and giving a rationale for procedures ■ The coaching function: making culturally avoided aspects of an illness approachable and understandable
Monitoring and ensuring the quality of healthcare practices	■ Providing a back-up system to ensure safe medical and nursing care ■ Assessing what can be safely omitted from or added to medical orders ■ Recognition of a generic recurring event or problem that requires a policy change ■ Seeking appropriate and timely responses from physicians
Organizational and work-role competencies	■ Coordinating, prioritizing, and meeting multiple patient needs and requests ■ Building and maintaining a therapeutic team to provide optimum therapy; providing emotional and situational support to nursing staff ■ Competencies developed to cope with staff and organizational resistance to change; showing acceptance of staff persons to resist system change; using formal research findings to initiate and facilitate system change; using mandated change to facilitate other changes ■ Making the bureaucracy respond to the patient's and family's needs
The consulting role	■ Providing patient care consultation to the nursing staff through direct patient intervention and follow-up ■ Interpreting the role of nursing in specific, clinical patient care situations to nursing and other professional staff ■ Providing patient advocacy by sensitizing staff to the dilemmas faced by patients and families seeking healthcare

Adapted from Spross and Baggerly (1989)

those that put the patient and/or organization at risk if not performed to standard. Critical competencies are ones that are critical for effective performance. Competencies required in the following key areas should be assessed to ensure that healthcare workers are able to perform infrequent, high-risk, and critical healthcare activities: performing duties, procedures, treatments, etc.; using equipment; emergency response and lifesaving interventions; managing patient/customer relations; patient assessment; and communicating with patient and/or family. Competency-related data may be derived from specialized tests, interviews, performance evaluations, quality improvement findings, patient satisfaction surveys, employee surveys, and other needs assessments. Based on the results of competency assessments, appropriate educational or remedial programs should be developed to meet identified needs or gaps and to improve overall performance.

How are competencies acquired?

As noted above, competence is defined in the context of knowledge, skills, abilities, and traits. These components of competence are acquired in different ways.

A provider obtains knowledge in several ways, including pre-service education and in-service training. Knowledge is further enhanced through on-the-job experience—including feedback from supervisors and peers—and continuing education. Field experience shows that providers may not use their knowledge correctly or consistently all the time for a variety of reasons (CHR 1999). Factors at the individual, organizational, and environmental levels all affect the correct use of knowledge.

Skills refers to "actions (and reactions) that an individual performs in a competent way in order to achieve a goal" (Ericsson 1996). Skills are gained through hands-on training using anatomic models or real patients, or through role-plays. One may have no skill, some skill, or complete mastery; therefore, in teaching or testing a skill, the level of acceptable mastery must be defined based on the training level.

Abilities refers to the power or capacity to do something or act physically, mentally, legally, morally, etc. Abilities are gained or developed over time and, as a result, are more stable than knowledge and skills. Traits influence abilities (discussed below).

Traits refers to distinguishing characteristics or qualities, especially of a personal nature. These include attitudes (personal and social values), self-control, and self-confidence. Traits influence abilities. For example, self-efficacy is the belief that one can do a task as required; it influences

whether a behavior will be initiated and sustained (Bandura 1986). Self-efficacy is determined by the confidence and/or training of a health worker. Low self-efficacy can lead to poor compliance with clinical guidelines and other standards of care. Many traits are slow to change or even permanent.
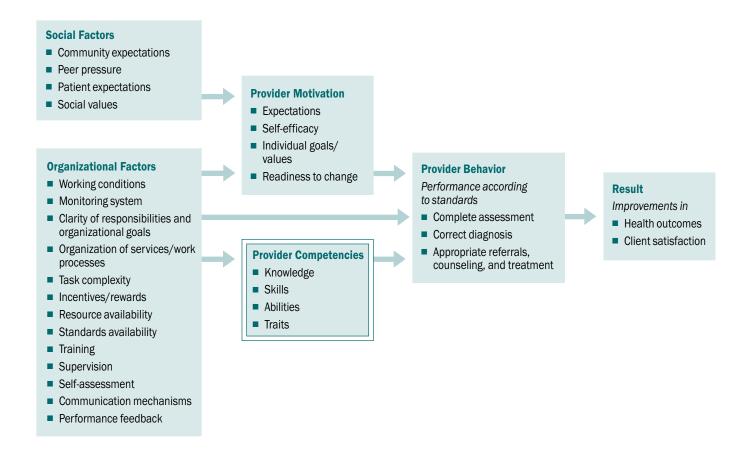
A prerequisite to provider behavior change is to understand the underpinnings of current practices. A health provider may not be able to overcome the persistence of previous practice or may not have the motivation to change (Cabana 1999). According to the "readiness for change" model, developed by Prochaska and DiClemente (1986), behavior change consists of a continuum of steps that include pre-contemplation, contemplation, preparation, action, and maintenance. For example, when this model was applied to physician attitudes towards cancer screening guidelines, the results suggested that nearly half of the physicians surveyed were in a pre-contemplation stage and not ready to change behavior (Main et al. 1995).

Other factors affecting provider performance

Extensive research shows that the more the person's competencies match the requirements of a job, the more effective the person will be performing (Hunter 1983; Spencer et al. 1994). However, competency does not always lead to effective performance. There is a difference between what an individual should be able to do at an expected level of achievement and what he or she actually does in a real-life setting (While 1994, p. 526). A number of other factors including personal motivation, adequate support of the hospital authorities, colleagues, and even non-professional health workers can affect worker performance (Salazar-Lindo et al. 1991). According to Campbell (1993), motivation is reflected in the completeness, the intensity, and the persistence of effort. For example, a healthcare worker may be competent to perform a medical procedure, but may not be willing to expend the effort to perform *all* the required behaviors. Figure 1 illustrates the relationship of these factors and their effect on provider performance.

Motivation is strengthened and providers work with more completeness, intensity, or persistence when: they are committed to a clear and challenging goal (Locke and Latham 1990), their job offers an opportunity to demonstrate mastery rather than an occasion to be evaluated (Nicholls 1984; Nicholls and Miller 1984), they believe that a particular procedure or standard will be effective (Cabana 1999), and they have high expectations for success ("outcome expectancy") (Bandura 1982, 1986; Bandura and Cervone 1986; Erez 1990). Individual traits can also determine motivation. For example, people with a disposition for accomplishing

Social Factors
- Community expectations
- Peer pressure
- Patient expectations
- Social values

Organizational Factors
- Working conditions
- Monitoring system
- Clarity of responsibilities and organizational goals
- Organization of services/work processes
- Task complexity
- Incentives/rewards
- Resource availability
- Standards availability
- Training
- Supervision
- Self-assessment
- Communication mechanisms
- Performance feedback

Provider Motivation
- Expectations
- Self-efficacy
- Individual goals/values
- Readiness to change

Provider Competencies
- Knowledge
- Skills
- Abilities
- Traits

Provider Behavior
Performance according to standards
- Complete assessment
- Correct diagnosis
- Appropriate referrals, counseling, and treatment

Result
Improvements in
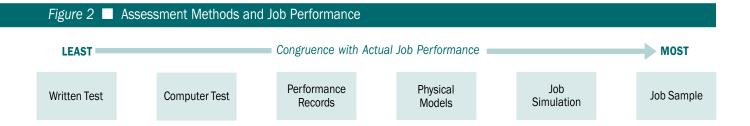- Health outcomes
- Client satisfaction

challenging objectives work harder (Spencer et al. 1994). For a thorough discussion of health worker motivation, see Franco et al. (2000).

Factors external to the individual and related to organizational and social conditions also influence provider behavior and performance. There is evidence that higher performance is associated with sufficient resources to perform the job, clear role expectations and standards of performance (Pritchard 1990), feedback on performance (Bartlem and Locke 1981; Locke and Latham 1990), and rewards that are contingent on good performance (Locke and Latham 1990) and the nature of the reward. In general, the expectations of the organization, profession, and community may influence the behavior and performance of providers for better or worse. For example, lack of supervision may result in some health workers' cutting corners inappropriately. Of course, unavailability of basic resources (such as equipment, supplies, and medicines) can result in poor performance in spite of high competence and motivation.

Provider performance varies and is partly determined by the social standing, awareness, and expectations of the client. Poorer, less educated, and less demanding clients often receive less attention (Schuler et al. 1985). Peer pressure plays a role, and some providers fail to comply with standards even if they have the requisite knowledge and skills.

The socio-cultural environment in which people were raised, live, and work affects job performance. "Any meaningful analysis of work motivation in developing societies has to be juxtaposed with an analysis of the physical and socio-cultural environment as well as the stable attributes of the individual who is a product of such an environment." (Mira and Kanungo 1994, p. 33). Therefore, organizations need to adopt management practices that are consistent with local conditions (Kanungo and Jaeger 1990). Mendonca and Kanungo (1994) propose that managers set goals within the employee's current competencies and increase those goals as the employee experiences success and feels more capable of achieving more difficult goals.

Figure 2 ■ Assessment Methods and Job Performance

LEAST → *Congruence with Actual Job Performance* → **MOST**

Written Test	Computer Test	Performance Records	Physical Models	Job Simulation	Job Sample

Approaches to competence measurement in healthcare

The term *gold standard* is used in healthcare (and other arenas) to describe practices that are accepted as the best and most effective for a particular problem, disease, or intervention. Although there are best practices that a healthcare provider should use, there is no gold standard for measuring the provider's competence. The most effective and feasible approach depends on the situation. This section presents different approaches for measuring competence.

Which assessment method should be used?

Selecting a measure of competence involves making three decisions: (a) What assessment methods are available? (b) How should scores be derived based on the method? and (c) Who should observe/evaluate competence?

a. Assessment methods

Competence can be measured using a variety of methods: written tests, computerized tests, records of performance, simulations with anatomic models, other simulations, job samples, and supervisory performance appraisals. These methods differ in a number of ways. Of particular relevance to this paper is the degree to which a method assesses a single competency (as compared to the assessment of multiple competencies) and the extent to which what is being observed or evaluated approximates on-the-job performance. The availability of resources is a key factor in selecting a particular method or methods and setting assessment frequency.

The advantage of assessing a single competency is that it could then be targeted for improvement. Competency measurements that predict job performance may increase the chances that a remedial action will be identified and will increase job performance of trainees who score low on end-of-training competency tests. Competency measurement methods that poorly predict job performance are less likely to be effective in this regard.

Assessment methods can be presented on a continuum reflecting the extent to which they approximate actual job performance, from least to most similar, as shown in Figure 2. Spencer et al. (1994) reported that job samples and job simulations are among the best predictors of job performance. Written tests are probably furthest from—and the weakest predictor of—actual job performance.

The greater the congruence between the method used to measure competency and the actual job, the more likely the competency measure will predict job performance, although this does not imply that good predictors of job performance are the best estimators of competency. In fact, competency cannot always be inferred from job performance. *Poor* job performance may have many causes, not just lack of competency, but *good* job performance usually does imply the presence of competencies needed for the job, especially for tasks that are reasonably complex. Simple tasks, however, can be performed adequately for a while with inadequate knowledge by emulating someone else's behavior. In a pre-post, control group research design, increased competency can be inferred when experimental group performance improves significantly more than control group performance following an intervention aimed at increasing competency.

Different assessment methods have different strengths and weaknesses. Both computerized and written tests can assess abilities, traits, and knowledge, but they cannot assess skills, which require the physical performance of some actions. Records of performance, unlike all the other methods for assessing competence, can be evaluated without the provider's awareness, but often omit important information about performance. Anatomic models limit both the types of competencies that can be assessed and the realism of the patient-provider interaction. Job samples and job simulations limit the circumstances and illnesses (or needs) and require both observation and evaluation. Fear about being evaluated can cause poorer performance, and awareness of being evaluated can contribute to greater care and attention to a patient, which may only appear to be better performance.

	Assessment Method							
	Job Sample	**Job Simulation**	**Anatomic Model**	**Clinical Simulation Testing**	**Records**	**Computerized Test**	**Written Test**	**Performance Appraisal**
Advantages								
Approximates real situation	X	X		X				
Assesses single or multiple competencies	X	X	X	X		X	X	X
Patient can report on care	X	X						
Replicable		X	X	X	X	X	X	X
Evaluates full range of competencies	X	X						X
Disadvantages								
Must wait for situation	X							
Requires extensive resources		X		X		X		
Requires trained assessor	X	X	X		X			
Potential bias								X

Table 2 summarizes the advantages and disadvantages of various competence measurement methods. Table 5 (see Appendix 1) provides the results of various studies that tested approaches for measuring provider competence. The following summarizes the results of healthcare research in each method.

Written tests

The patient vignette is a type of written competency test in which short case histories are presented, and test-takers are asked pertinent questions about what actions should be taken if the portrayal were real. Patient vignettes measure competency in applying knowledge. Advantages include standardization of questions, objectivity in scoring, and minimal costs. One disadvantage is that competencies involving physical skills, traits, and abilities cannot be measured. In addition, performance on tests is inconsistently predictive of performance with patients (Jansen et al. 1995; Newble and Swanson 1988, Van der Vleuten and Swanson 1990).

Computerized tests

One of the disadvantages of measuring competence with patients or models is that a trained assessor is needed. An alternative, being used in the nursing profession and which can be used to assess clinical decision-making skills, is computerized clinical simulation testing (CST). Although many computer tests essentially replicate a paper test, recent technological developments enable the creation of tests that more closely approximate actual job conditions. "In CST, examinees are not cued to patient problems or possible courses of action by the presentation of questions with decision options. Instead, a brief introduction is presented and the desired nursing actions are then specified by the examinee through 'free text' entry using the computer keyboard." (Bersky and Yocom 1994) The first screen presents the case, and then the test-taker requests patient data using free text entry. The program responds by searching a database of 15,000 nursing activity terms, such as vital signs, progress notes, or lab results. Realism is enhanced by having the patient's condition change in response to the implementation of nursing interventions, medical orders, and the natural course of the underlying health problem. The actions of the test-taker are compared to standards for a minimally competent, beginning-level nurse. Major advantages of CST (in addition to those in Table 2) are consistency of the cases, objectivity in scoring, and low cost once the program is developed. Drawbacks include the inability to evaluate competencies involving physical actions and interpersonal interactions, high development costs, and lack of computers in many developing countries.

Review of medical records

Medical records have some advantages and many disadvantages as a data source for assessing competency. On the positive side, data can be obtained retroactively and at a relatively low cost compared to other methods. Providers are not aware of or influenced by the data collection, eliminating one source of bias. On the negative side, patient records are often incomplete. Using standardized patients, Norman et al. (1985) analyzed the completeness of patient records and found that many omitted critical actions. For instance, counseling was rarely recorded. Overall one-third to one-half of the procedures performed were not recorded. Furthermore, record audits proved unlikely to detect missing diagnoses or misdiagnoses. Missing and poor quality records are prevalent in developing countries, especially in primary care facilities. Another problem is that competency cannot be reliably inferred from performance.

Anatomic models

Anatomic models are often used in healthcare for training and competency measurement. They are especially appropriate for assessing competency (as opposed to performance) in certain physical skills, such as inserting an intravenous needle (IV) or an intrauterine device (IUD), or cardio-pulmonary resuscitation (CPR). Other advantages include low cost, standardized testing, and repeated use without burdening patients. The disadvantages include their inability to simulate: (a) provider-client interactions (including client feedback), and (b) the complications that occur in real patients, such as multiple or inconsistent symptoms.

Job simulation

In actual practice, providers work with real clients in real job settings. In a job simulation, the clients, the setting, or both are not real. In the last 20 years, two job simulation techniques—*standardized clients* (either announced or unannounced) and the *objective structured clinical examination*—have emerged as simulation methods for assessing provider competence.

Standardized clients can be either real clients or healthy individuals who have been trained to provide a reproducible and unbiased presentation of an actual patient case (Tamblyn et al. 1991). Further, standardized clients can be either *announced* or *unannounced*. With announced standardized clients, the provider is made aware that the client is standardized and is a healthy individual pretending to have a medical concern. With unannounced standardized clients (sometimes referred to as "mystery clients"), the provider is not informed that the client has been trained to perform as a patient. Studies show that experienced physicians cannot differentiate real patients from unannounced

standardized patients and that history taking, physical examination, findings, and diagnoses are quite similar for announced and unannounced standardized patients.

One advantage of standardized clients is that they can be trained to accurately and consistently evaluate and report provider performance. In one study (Colliver and Williams 1993), standardized patients consistently agreed with 83 percent of the evaluations of clinical skills made by three faculty physician observers. In another, standardized patients were 95 percent accurate in portraying the details of an illness (Colliver and Williams 1993). Standardized clients also provide a replicable case for multiple healthcare providers, thus enabling direct comparison of their performances (Stillman et al. 1986). In addition, unlike an actual client, a standardized client can portray the disease or problem in a way that is most relevant to the particular competency being measured.

Use of standardized clients also has disadvantages. It can be difficult to separate competence from performance. Studies show that clinical performance by a single provider is not consistent across patients and specifics of a disease (Stillman et al. 1991); therefore, one standardized patient does not provide a reliable estimate of provider performance or competence. Furthermore, unlike real clients, standardized clients must be paid and trained (Stillman 1993). Tamblyn et al. (1991) reported that it took three one-hour training sessions to train standardized patients and that more experienced standardized patients performed more accurately than less experienced ones. If the competence of several healthcare providers is being assessed, standardized clients must be willing to present their stories several times. Certain symptoms cannot be simulated in healthy individuals and require the use of real clients as standardized clients (Tamblyn et al. 1991). This method also poses challenges if providers' surgical competencies (e.g., minilaparotomy under local anesthesia) are to be assessed.

The *objective structured clinical examination (OSCE)* measures clinical skills using a uniform, structured format of rotating stations. A healthcare provider performs a clinical task at each station. OSCE assesses a variety of competencies and has been most often used in medical schools and residency programs. It has the same advantages and disadvantages as standardized clients except that the latter makes greater demands of time and resources, such as examining rooms, trained examiners, preparation of materials, training of and payments to standardized patients, equipment and materials (Cusimano et al. 1994). Research indicates that three to four hours (per examinee) are necessary for the consistent measurement of performance. Time spent per station does not appear to be an important factor

Table 3 ■ Advantages and Disadvantages of Scoring Methods

	Scoring Methods		
	Checklists	Rating Scales	Overall Assessments
Advantages			
Can be used with limited training	X		
Useful for self-evaluation	X	X	
Disadvantages			
Often time-consuming	X		
Differences in interpretation		X	X
Requires training to use		X	X
Requires expertise to use			X
May leave out important information			X

in the quality of the measurement (Elnicki et al. 1993). Medical students who were evaluated using this method thought that the OSCE was fair and clinically relevant (McFaul and Howie 1993).

Job sample

Competence is sometimes inferred from measurements of provider performance with a sample of real patients in an actual job setting. There are, however, difficulties in measuring job performance, including the unpredictability of the environment, waiting for a clinical situation to arise to accommodate testing (Ready 1994), and differences in the levels of difficulty among different cases. In addition, a single observation of performance does not provide a reliable estimate of the provider's competence. Studies of clinicians' performances indicate that the quality of service varies widely from one patient to another (Stillman et al. 1986). In a specific area of practice, the provider interacts with patients who differ in terms of illness or need, works under a variety of circumstances, performs activities requiring a range of competencies, and proceeds without formal observation or evaluation. Multiple observations of cases with somewhat similar problems would provide better information about a provider's competencies.

Performance appraisals

Periodic appraisals by supervisors or peers and self-assessments can also be used to infer competence. Supervisor and peer appraisals can use data from multiple sources, including observations of provider-patient interactions, record reviews, patient interviews, self-appraisals, and sentinel event data to assess competence. Using multiple data sources helps to reduce assessor bias. Self-assessments can use pre-structured checklists to reduce bias when identifying areas of poor performance (Bose et al. forthcoming). Performance appraisals suffer from the difficulties of inferring competence from performance.

b. How should scores be derived?

Regardless of the method used for measuring competence, a scoring technique is needed. In order to develop a measure, data (test scores, observations, records) need to be analyzed to derive scores, which indicate the provider's level or extent of competence. The analysis involves comparing the data to a standard of competence.

A level of competence is defined in relation to a standard. Competence can be above or below the standard, or in some cases it can be stated as a proportion of the standard, such as 70 percent of the standard. Competency standards answer the question, "At what level of performance is someone considered competent, so that we can trust him or her with the healthcare of patients?" Extensive research shows that raters evaluate performance differently, so a lack of standards is likely to yield inconsistent evaluations of competence (Fitzpatrick et al. 1994; Harden and Gleeson 1979; Stillman et al. 1991; Van der Vleuten et al. 1991).

How are standards determined? The most objective standards are usually based on scientific evidence and/or expert consensus (Marquez forthcoming). It is difficult to argue against a standard of performance with known links to positive patient outcomes (Benner 1982). International and national health organizations such as the World Health Organization, medical societies, the Agency for Healthcare Research and Quality, and national health ministries establish standards for patient care based on evidence from the scientific literature as interpreted by expert panels. Groups of professors often identify standards of performance for evaluating the competence of medical or nursing students. Standards of competence can also be based on the performance of individuals selected as excellent clinicians (Sloan et al. 1993). Alternatively, behaviors that differentiate between experienced and inexperienced providers can be used as the standards of competent performance (Benner 1982).

Table 4 ■ Advantages and Disadvantages of Different Types of Assessors of Competence

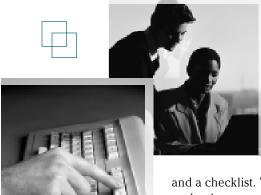|	Type of Assessor						
	Self	Trainer	Peer	Supervisor	Untrained Patients	Trained Patients	Expert Observer
Advantages							
Can observe competence or performance	X	X	X	X	X	X	X
Can evaluate competence or performance		X	X	X		X	X
Has many opportunities to observe performance	X		X	X			
Disadvantages							
Bias due to halo effect		X	X	X			
Provokes anxiety		X		X			X

The three methods most often used to attach a score to the level of competence are checklists, rating scales, and overall assessments. Checklists provide a pre-defined list of behaviors (the standards) that are judged to be at or below standard for a particular provider. The final score is often simply the number (or percentage) of behaviors performed at standard. Rating scales provide a range of possible responses for each behavior on the checklist; the scales reflect the level of competence or performance attained with respect to that behavior. For example, the level could range from 1 (worst) to 5 (best). The precision of the definition of the different levels varies widely. The individual behavior scores are summed to obtain an overall score. Overall assessments rely on the general evaluation of the rater and exclude explicit evaluations of individual behaviors. Table 3 summarizes the advantages and disadvantages of these scoring methods.

Checklists are the least subjective of the three scoring methods, and overall assessments the most. Rating scales tend to be more subjective than checklists because the levels for each behavior are rarely defined as precisely as the behaviors in the checklist. The historical trend in scoring methods is towards objectivity, to use methods that reduce judgment and increase agreement, which means in the direction of detailed checklists and away from overall evaluations. Research findings show greatest agreement among independent raters who use checklists rather than the other two methods (Van der Vleuten et al. 1991).

Detailed checklists are particularly appropriate for use by less expert and less well-trained personnel, such as standardized patients (Norman et al. 1991). Checklists are also useful for training and self-evaluations, since they clearly define the steps involved in performing a task. However, they may be less appropriate for the assessment of complex competencies (such as interpersonal skills or a complex clinical/surgical procedure) that can be difficult to describe in terms of specific, discrete behaviors. Van der Vleuten et al. (1991) reported that checklists in an OSCE ranged from 30 to 120 items long. Observing, recording, and checking through long lists of behaviors is an arduous task and may exceed the willingness of observers.

Like checklists, rating scales focus the observer's attention on those aspects of competence that are important and provide a convenient way to record judgments (Fitzpatrick et al. 1994). One problem with rating scales is that they yield inconsistent measurements if they lack precise definitions for each rating level, allow subjective interpretations about the performance required for each level, or are unclear about the meaning of terms that define the levels.

Overall assessments require greater expertise than either rating scales or checklists since the assessor needs to know what to observe and evaluate, as well as the standards of competence. A significant disadvantage of overall assessments is that the assessor does not necessarily list all of his or her observations. Hebers et al. (1989) had physicians observe a videotape of a resident's performance and score the performance using overall evaluations, a rating scale,

and a checklist. The overall evaluations were not specific and included few comments on good performance. Comparison of the ratings and the detailed checklists with the overall evaluations indicated that although the assessors identified specific examples of good and poor performance, they did not include many of these examples on the general comment sheet.

c. Who should assess competence?

Information about a healthcare worker's competence can be provided by different types of people, including colleagues (e.g., supervisors, peers, trainers, professors), patients, independent observers, and the provider him- or herself. Table 4 summarizes advantages and disadvantages of these different types of raters. Research on performance and competency appraisal does not provide evidence about the relative superiority of one type over another (Borman 1974; Klimoski and London 1974). The use of multi-source performance appraisals (e.g., 360-degree feedback) suggests that evaluative information from many sources provides a more complete picture of performance than evaluations from only one perspective (Franco et al. 2000).

Assessors are not interchangeable. Research on the evaluation of job performance indicates that assessors who occupy different positions in an organization notice different aspects of performance and, thus, evaluate performance differently (Mohrman et al. 1989). For example, peer appraisals focus on the way employees relate to each other (Latham 1986). When peers evaluate a colleague, they tend to compare the colleague's performance with that of co-workers or with their own performance (Mumford 1983). In some environments, peers may be unwilling to appraise each other because they view evaluation as a management job and assume that their role is to protect (i.e., provide no negative information about) their colleagues.

Supervisors and peers have many opportunities to observe providers and generally know them well. As a result, they may be able to give a more accurate and thorough evaluation of competence than either patients or independent observers. However, it may be difficult for someone who knows the healthcare provider to make an unbiased assessment of specific competencies, independent from their overall assessment of the person (sometimes referred to as the *halo effect*[1]).

Not only do different types of assessors rate differently; the people being assessed respond differently. Use of supervisors to assess competence increases the perception that the testing environment is threatening. This anxiety can contribute to lower performance during testing (Ready 1994). Mohrman et al. (1989) reported that the act of measurement may influence performance or competence under test conditions, either degrading it (by increasing anxiety) or improving it (because the person under assessment is making his or her best effort).

Regardless of who is selected to rate performance, it is essential that the assessor understand the standards for effective performance. The assessor can either be an expert on the topic or can be trained to observe and evaluate specific competencies. Research indicates that a trained, but non-expert, assessor can provide accurate assessments of competence. Standardized patients, given training by medical school faculty, reliably evaluated the same clinical skills as those evaluated by faculty physicians (Colliver and Williams 1993). MacDonald (1995) reported that trained midwives were as accurate in their evaluation of peers as the gold standard observers. Even untrained people, such as mothers who observe interactions between health workers and their children, can accurately report what the workers do (Hermida et al. 1996). In this latter situation, someone who understands competence can use these behavioral reports to prepare an evaluation of competence.

Criteria for selecting measurement methods

Validity

Validity concerns the degree to which a particular measurement actually measures what it purports to measure. Validity provides a direct check on how well the measure fulfills its function (Anastasi 1976). Does it truly measure the particular competency it intends to? Is it capturing a different competency? Or, more frequent, is it measuring performance rather than competency? Does it identify gaps in knowledge, skills, abilities, or traits that are needed for

---

[1] The halo effect is the generalization from the perception of one outstanding personality trait to an overly favorable evaluation of the whole personality.

the job and that can be corrected through training, experience, or better methods for matching individuals to jobs?

Below are some findings from the literature about the validity of various competency measures:

- Assessments of medical records are not good indicators of healthcare provider competence, largely because many medical procedures are not recorded (Franco et al. 1997; Hermida et al. 1996; Norman et al. 1985)

- Performance on tests is inconsistently correlated with performance with patients (Jansen et al. 1995; Sloan et al. 1993). Written, oral, and computerized tests are primarily measures of knowledge: patient care requires several skills and abilities in addition to knowledge

- There is substantial evidence that the OSCE method can be used to effectively assess a wide variety of competencies (Colliver and Williams 1993; Elnicki et al. 1993; Stillman et al. 1986)

- Norman et al. (1985) found only moderate correlation between the evaluations of performance by the standardized-patient method and by patient records. Lyons (1974) reported similar low correlations between medical records and medical care

Reliability

Reliability refers to the consistency of scores for a particular person with respect to a particular competency when evaluated by different methods, by different raters, or for more than one patient. Below are some of the findings about the reliability of competency measures.

- Raters use different criteria in evaluating the competence of healthcare providers (Norman et al. 1985). The consistence in ratings provided by different raters is low to moderate (Stillman 1993)

- Performance with one patient does not represent performance on other cases (Cohen et al. 1996; Franco et al. 1996; Norman et al. 1985; Reznick et al. 1992). In order to obtain adequate reliability, multiple patients are required (Colliver and Williams 1993; Sloan et al. 1993)

- Clarifying the checklist improves the reliability of the ratings (Colliver and Williams, 1993)

- Trained peers can accurately assess healthcare provider performance (MacDonald 1995)

- Healthcare provider self-assessments are consistent with their observed performance with patients (Jansen et al. 1995; MacDonald 1995; Bose et al. forthcoming)

- Both patients and caretakers of patients accurately report healthcare provider performance (Colliver and Williams 1993; Franco et al. 1996; Hermida et al. 1996). While these reports can be used as inputs into an evaluation of competence, they are not measures of competence

Feasibility

Resources are required to design and implement an assessment of competence. Decisions to determine which measure to use should reflect the following issues:

- The number of individuals to be assessed

- The time available for the assessment

- The willingness of the assessor to use the assessment instrument

- The willingness of the healthcare provider to accept the assessment

- The extent of training available those who will participate in the assessment

- The resources (funding, assessors, equipment, space, etc.) available for the assessment. For example, OSCE procedures require extensive resources and may be cost-effective only when evaluating a number of individuals on a variety of competencies

- The time, staff, and funding available for development and pretesting of instruments. Computerized simulations are expensive but offer promise for evaluating knowledge and decision-making skill

- The competency to be assessed. For example, competency in removing Norplant implants is much easier to measure than competency in managing complications in a delivery

Research and implementation needs

According to discussions with health program managers and international health experts, healthcare provider competencies should be measured periodically. Limited information is available on competency measurement in developing country health programs. This section raises other needs.

What is the relationship between competence and performance? There is extensive evidence that: (a) although competency affects performance, the relationship is not direct, and (b) other factors (the work setting, time, and motivation) play a major role in determining performance. More research is needed to study the relationship between

Lecture programs and conferences disseminate information about new innovations among health workers and other staff. These programs cover information on the current scope of practice or changes in the art and science, based upon scientific information learned from current medical research.

Continuing education (CE) courses are an effective way of keeping health workers abreast of new innovations in their field of specialization. In some countries, health workers must earn a minimum number of CE units every year as part of re-certification.

Refresher programs enable health workers to review the original training program in a condensed number of hours. However, refresher programs do not help in expanding the cognitive or psychomotor ability above the entry level.

Self-education is another way health providers can improve knowledge in specific areas. This method can be useful for acquiring new knowledge or skills with immediate application to a task. Self-education may involve manuals or computer-based training (CBT: Knebel 2000).

Case reviews rely on patient-care reports, audio/video tapes of services, and laboratory and other diagnostic reports. Case reviews provide an opportunity for discussion, usually led by a mentor or preceptor, and appear to be more effective than other didactic methods in improving knowledge or skills.

Grand rounds at health facilities with a diverse patient population provide a unique learning opportunity for novices. Through grand rounds, novices are exposed to the wider continuum of patient care, and they can interact with other members of the healthcare team and experience a higher level of cognition about patient conditions and recovery than would normally occur outside the hospital setting.

Sentinel-event review, involving a review of an unexpected occurrence (mortality or morbidity), provides an opportunity to conduct a root cause analysis. Results should form the basis of a plan to reduce risk. The plan must be monitored to evaluate its effectiveness relative to the root cause.

Mentoring/precepting, provided by more experienced health professionals, is a good strategy for improving skills of novices and new-entry health professionals.

---

competence, performance, and these other factors, and to analyze the interaction between competence and these factors on performance. In addition, research should be undertaken to provide guidance to organizations that are planning to evaluate competence to determine feasible strategies to effectively assess competencies.

How can improvements in competencies achieved by training be sustained? Most training programs demonstrate improvements in immediate post-training competence levels. However, various studies show these improvements decay over time, sometimes rapidly. Little research exists on the most cost-effective ways to maintain the improvements gained during training. This should be remedied (Kim et al. 2000).

Which competency measures (detailed checklists, rating scales, overall assessments, simulations, work samples) are most relevant for assessing the competence of various levels of healthcare workers (community-based workers, paramedics, nurses, doctors) in developing countries? Limited research has been conducted on identifying cost-effective approaches for measuring provider competence in developing countries. Although, a lot of checklists and questionnaires have been developed and used for assessing provider knowledge, limited models for assessing provider skills have been designed (but see Kelley et al. 2000).

Who is best suited to conduct competency assessments in developing countries? What types of measures would they be willing to use?

What measures would healthcare workers understand, accept, and use in developing countries?

Research is needed to identify differences in assessment results when using different assessors (supervisors, peers, external reviewers, etc.).

Should competency measurement be made part of licensure, certification, and accreditation requirements?

Quizzes and questionnaires are excellent tools for assessing provider knowledge. Results can be used to develop remedial strategies for improving provider knowledge.

Practical performance examinations or observations with patients (real or simulated) are cost-effective ways to assess providers' psychomotor and interpersonal skills. These methods can be administered in a minimal amount of time and cover a wide domain of practice. Exam results can be quickly tabulated for making timely decisions about remedial actions.

Review of records and/or sentinel events is also cost-effective and measures competency. A root cause analysis of a sentinel event can shed light on a provider's rationale in following a specific treatment regimen.

Supervised patient interactions also measure competence. A health provider can be observed while performing skills and procedures on a diverse patient population in a relatively short period.

Hospital clinical performance evaluations can be used to identify performance problems and their root causes. Based on this analysis, competence of providers in providing quality care can be assessed.

Developing country health programs should consider making periodic competency assessments as part of the licensure and accreditation requirements. However, there is a concern that giving too much control to government institutions for licensure and accreditation will complicate issues and create bottlenecks. Research is needed to identify appropriate public-private mechanisms for licensure and re-certification, where individual and organizational competencies are taken into account.

How can feedback from competency assessments be used to improve compliance by health providers? Feedback about competence and performance can motivate changes in provider attitudes and behaviors in developing countries. Research is needed to identify cost-effective approaches to provider feedback (one-to-one, one-to-group, newsletters, etc.).

Does an understanding of competency requirements for a specific job improve the selection and training of staff?

Health staff is often recruited without conducting an in-depth analysis of competency requirements. There is a growing belief that it is simply impossible to improve a person's performance until specific competencies required for satisfactory or superior performance are identified. It is only after these competencies are identified that staff can be selected. Some new recruits may require skill or knowledge enhancement to be able to perform optimally. Research is needed to identify the impact of clearly defined competency requirements on staff performance, if any.

Conclusion

The literature suggests several conclusions, summarized below, concerning the measurement of competency:

Competency can be assessed using tests or inferred from performance that has been assessed using simulations or work samples. The major advantage of tests is that single competencies can be distinguished and targeted for improvement. The major advantage of simulated patients and job samples is that they are more predictive of job performance.

Competency is not performance. Although competency can predict performance, a competent healthcare provider may not necessarily use effective procedures on the job. Both internal factors (motivation, agreement with a standard, self-efficacy, inertia, etc.) and external factors (supervision, feedback, availability of resources, community, peer expectations, and incentives) affect whether a healthcare provider will apply his or her competency.

Detailed and immediate feedback to the healthcare provider about his or her competence is useful for both learning and improving performance.

Standards of competency must be defined carefully, even for expert assessors. Clear statements about both the indicators and the levels of competency improve the reliability and validity of the assessment.

Competency can be measured using a variety of methods, including tests presented in a written format or on a computer, in an interview, or through a simulation or work sample. Each assessment method has its strengths and weaknesses. While written tests or interviews can assess knowledge, assessments using models or job samples can closely assess skills.

All assessors must be trained to give accurate reports and evaluations of competency. The length and content of the training depends on the expertise of the assessor, the competency to be assessed, the assessment instrument used, and the conditions for evaluation.

Different types of assessors, including supervisors, peers, patients, and observers, can accurately report and assess healthcare provider competency.

Checklists are particularly useful for giving feedback to the healthcare provider and for use by non-expert raters. Detailed checklists are also useful for self-evaluation as they provide a teaching tool and a job aid. However, some raters may be unwilling to use checklists, which can be long and time-consuming. It is necessary to determine a length that provides useful feedback and that assessors will use.

Written or computerized tests are an effective way to measure knowledge but not to assess skills that are required for some tasks.

The effective ways to measure competency include evaluations of performance by experts or trained observers, reports from patients (especially trained or standardized patients), and objective structured clinical examinations.

The literature review also highlights the need to promote competency measures in developing country health programs. Most developing country health programs do not use measures to identify competency gaps. The introduction of periodic competency assessments will assist in improving the quality of care.

Appendix

Table 5 ■ Selected Studies on Competence Measurement of Healthcare Providers					
Author	**Target Group**	**Target Area** (Focus of Study)	**Description of Intervention**	**Statistically Significant Improvement**	**Other Findings**
Das et al. (1998)	Medical students (UAE)	Comparison of self- and tutor evaluation results	First-year medical students (64) undergoing problem-based learning (PBL) underwent self-assessments as well as tutor evaluations.	Self-evaluation results and tutor assessment scores were similar, but male student self-evaluation scores were higher than female students' on overall scores.	The sharing of assessment reports between students and tutors was perceived to be useful for the students' development of the skills of analysis, differentiation, and critical appraisal.
Devitt and Palmer (1998)	Third-year medical students (USA)	Use of computers in assessing competence	History-taking and physical examination skills of 136 students were assessed in a series of structured and observed clinical stations and compared to similar computer-based problems.	Students scored equally on the computer-based tasks and in the observed stations, but the weaker students who passed on a clinic station were likely to fail on the computer task.	Study shows that computer-based clinical simulations can be constructed to supplement conventional assessment processes in clinical medicine and may have a role in increasing their reliability.
Elnicki et al. (1993)	Internal medicine junior clerkship (USA)	Reliability and content validity of objective structured clinical examination (OSCE) in assessing clinical competence	Third-year medical students (68), 6 residents, and 9 students finishing their medical rotation completed the 15 OSCE stations; each station took 15 minutes for patient examination.	Results showed that OSCE is a robust way of finding the competence of students, and the results were comparable to other methods of evaluation.	OSCE is not meant to evaluate cognitive achievement or interpersonal skills, so it is essential to use other methods of evaluation in conjunction with OSCE.
Friedman et al. (1999)	Physicians (USA)	Enhancement of clinicians' diagnostic reasoning by computer-based consultation	Two computer-based decision support systems (DSS) were used by 216 participants for diagnostic evaluation of 36 cases based on actual patients. After training, each subject evaluated 9 of the 36 cases, first without and then with a DSS, and suggested an ordered list of diagnostic hypotheses after each evaluation.	Correct diagnoses appeared in subjects' hypothesis lists for 39.5 percent of cases without DSS and 45.4 percent of cases after DSS. The study supports the idea that "hands-on" use of diagnostic DSS can influence diagnostic reasoning of clinicians.	
Hodges et al. (1999)	Medical staff	Effectiveness of binary content checklists in measuring increasing levels of clinical competence	Forty-two clinical clerks, family practice residents, and family physicians participated in two 15-minute standardized-patient interviews. An examiner rated each participant's performance using a binary content checklist and a global processing rating. The participants provided a diagnosis two minutes into and at the end of the interview.	Binary checklists proved to be less valid for measuring increasing clinical competence. On global scales, the experienced clinicians scored significantly better than did the residents and clerks, but on checklists, the experienced clinicians scored significantly worse than did the residents and clerks.	Diagnostic accuracy increased for all groups between the two-minute and 15-minute marks without significant differences between the groups.

Author	Target Group	Target Area (Focus of Study)	Description of Intervention	Statistically Significant Improvement	Other Findings
Hull et al. (1995)	Internal medicine clerks (USA)	Validity of different methods of assessing clinical competence	Multitrait-multimethod was used to assess student performance. Methods were clinical evaluation form (CEF), OSCE, and National Board of Medical Examiners.	The results suggest that there is a statistically significant but lower than expected convergence in the measures of clinical skills and knowledge across the three assessment methods.	
Jansen et al. (1995)	General practitioners (GP) and trainees in general practice	Use of performance-based test (PBT), a written knowledge test of skills (KTS), and a self-assessment questionnaire (SAQ) to assess provider competence	The three different assessment tools were administered to 49 GPs and 47 trainees in general practice.	The mean scores on the PBT and KTS showed no substantial differences between GPs and trainees, while GPs scored higher on the SAQ.	Performance-based testing is a better way to assess proficiency in hands-on skills.
McFaul and Howie (1993)	Medical students (UK)	Compare and assess clinical competence among final-year medical students of two medical schools	OSCE with stations designed to assess student competencies in history-taking, physical examination, interpretation of data or results, interpersonal skills, practical procedures, and factual knowledge	For a majority of stations, there was no statistically significant difference between the mean scores received by the students of two medical schools.	OSCE, well accepted by students and teachers, can be introduced easily in a medical school.
Peabody et al. (2000)	Staff physicians (USA)	Comparison of vignettes, standardized patients, and chart abstraction methods for measuring the quality of care	Staff physicians were randomly selected at 2 facilities. Performance was measured on 4 common outpatient conditions: low back pain, diabetes mellitus, chronic obstructive pulmonary disease, and coronary artery disease.	Using vignettes consistently produced scores closer to the gold standard of standardized patients than using chart abstractions. This pattern was found to be robust when the scores were disaggregated by case complexity, by site, and by level of physician training.	The study also found that low competence may be significantly determined by physician characteristics and not merely structural effects.
Ram et al. (1999)	Family physicians (USA)	Comparative study of measurement characteristics of a multiple-station examination (MSE) using standardized patients and a practice video assessment (PVA) of regular consultations in daily practice	Consultations of 90 family physicians were videotaped both in MSE and their daily practices. Peer observers used a validated instrument (MAAS-Global) to assess the physicians' communication with patients and their medical performances.	PVA was better able to assess performance of physicians' practices than MSE using standardized patients. Content validity of the PVA was superior to that of the MSE, since the domain of general family practice was better covered.	Observed participants judged the videotaped practice consultations to be "natural," whereas most physicians did not see their usual practice behavior while reviewing videotaped consultations of the MSE.
Ready (1994)	Nurses (USA)	Periodic clinical competency testing of emergency nurses as part of the credentialing process	All clinical nurses providing patient care were evaluated using cognitive and psychomotor tests. While cognitive skills were evaluated by written examinations, practical skills were assessed by following a sequential set of psychomotor skills.	The competency-testing program helped emergency nurses apply established standards of care more consistently.	

Author	Target Group	Target Area (Focus of Study)	Description of Intervention	Statistically Significant Improvement	Other Findings
Reznick et al. (1992)	First- and second-year residents at 4 sites (USA)	Validity of OSCE for assessing clinical competence of providers for the issuance of a license to practice medicine	The 240 residents were randomly assigned to take the 20-station OSCE test for national licensure.	The test results showed that OSCE is the state-of-the-art method for testing clinical skills as it provided a major advantage of observing actual performance. Also, the use of standardized patients enabled this examination to stimulate real-world conditions with a high degree of fidelity.	The study also showed that the production of this type of examination could not be rushed. Adequate time is needed for case development, standard setting, translation, and the technical formatting necessary for computer formatting. In addition, an examination of this scale requires that all aspects of development, production, and site supervision be centralized.
Sloan et al. (1993)	Surgical interns (USA)	Efficacy of OSCE in evaluating resident competence	Comprehensive 35-station OSCE was administered to 23 in-coming interns and 7 out-going interns.	The out-going interns performed significantly better than the in-coming ones.	
Snyder and Smit (1998)	Evaluators (Australia)	Level of rater reliability for evaluators of Emergency Medical Technicians practical examinations using the Michigan practical examination instrument	Licensed instructor-coordinators (104) were assessed on their scoring of two practical examinations on videotape, one passing and one failing performance.	Variations were found in scores given by evaluators for a single observed student. Also, evaluators often did not agree with a student's skill performance using the Michigan practical examination.	
Steiner et al. (1998)	Medical students (USA)	Effect of clinical training setting on skills	Medical students, who had encountered patients with low-back pain in primary care settings, tertiary care settings, both, or neither, were tested using standardized patients for their skills in the areas of history-taking, physical examination, and the selection of a diagnostic strategy.	Overall, students, irrespective of the training setting, performed poorly, suggesting that the curriculum inadequately teaches clinical skills needed to assess and manage common problems.	The training outcomes can be improved by setting clearer expectations of competencies and by setting mechanisms that assure that preceptors in ambulatory settings will help students meet those expectations.
Swartz et al. (1999)	Medical students (USA)	Comparison of validity of global ratings of checklists and observed clinical performance with standardized patients	Five faculty physicians independently observed and rated videotaped performances of 44 medical students on the seven standardized patient (SP) cases. A year later, the same panel of raters reviewed and rated checklists for the same 44 students on 5 of the same SP cases.	The mean global ratings of clinical competence were higher with videotapes than checklists, whereas the mean global ratings of interpersonal and communication skills were lower with videotapes. The results raise serious questions about the viability of global ratings of checklists as an alternative to ratings of observed clinical performance as a criterion for SP assessment.	

Table 5 ■ Selected Studies on Competence Measurement of Healthcare Providers *(Continued)*

Author	Target Group	Target Area (Focus of Study)	Description of Intervention	Statistically Significant Improvement	Other Findings
Wood and O'Donnell (2000)	Medical professionals	Measuring competence at job interviews: a traditional interview gives interviewers a chance to assess the intellect, enthusiasm, and "sparkle" of candidates, which cannot be faithfully conveyed in a curriculum vitae, but assesses competence and achievement poorly.	"Objective Assessment of Competence Achievement" (OACA) was designed to assess the competence of applicants. The system follows the OSCE model, whereby the interview panel split into several rooms, each assessing a different aspect of performance, e.g., consultation skills, achievements, and interpersonal relationships.	The OACA approach was better able to identify candidates with low competence. Most of the "poorly" competent candidates would have sailed through interviews.	More stations are needed for a comprehensive assessment. These stations could have reviewed prior assessments of skill or competence done at the workplace, such as videos of operations or a consultation.

References

Aguinis, H., and K. Kraiger. 1997. Practicing what we preach: Competency-based assessment of industrial/organizational psychology graduate students. *The Industrial-Organizational Psychologist* 34:34–40.

American Society for Training and Development. 1996. Linking training to performance goals. *Info-Line* (Issue 9606).

Anastasi, A. 1976. *Psychological Testing.* (4th ed.) New York: Macmillan.

Bandura, A. 1982. Self-efficacy mechanism in human agency. *American Psychologist* 37:122–47.

Bandura, A. 1986. *Social Foundations of Thought and Action: A Social Cognitive Theory.* Englewood Cliffs, NJ: Prentice Hall, Inc.

Bandura, A., and D. Cervone. 1986. Differential engagement of self-reactive influences in cognitive motivation. *Organizational Behavior and Human Decision Processes* 38:92–113.

Bartlem, C.S., and E.A. Locke. 1981. The Coch and French study: A critique and reinterpretation. *Human Relations* 34:555–66.

Bashook, P.G., and J. Parboosingh. 1998. Continuing medical education: Recertification and the maintenance of competence. *British Journal of Medicine* 316:545–48.

Benner, P. 1982. Issues in competency-based testing. *Nursing Outlook* 303–09.

Benner, P. 1984. *From Novice to Expert: Excellence and Power in Clinical Nursing Practice.* California: Addison-Wesley.

Beracochea, E., R. Dickson, P. Freeman, and J. Thomason. 1995. Case management quality assessment in rural areas of Papua, New Guinea. *Tropical Doctor* 25:69–74.

Berden, H.J.J.M., J.M.A. Hendrick, J.P.E.J. Van Dooren, F.F. Wilems, N.H.J. Pijls, and J.T.A. Knape. 1993. A comparison of resuscitation skills of qualified general nurses and ambulance nurses in the Netherlands. *Heart & Lung* 22:509–15.

Bersky, A.K., and C.J. Yocom. 1994. Computerized clinical simulation testing: Its use for competence assessment in nursing. *Nursing & Healthcare* 15:120–27.

Borman, W.C. 1974. The rating of individuals in organizations: An alternative approach. *Organizational Behavior and Human Performance* 12:105–24.

Bose, S., E. Oliveras, and W.N. Edson. Forthcoming. How can self-assessment improve the quality of healthcare? *Operations Research Issue Paper.* Bethesda, MD: To be published for the U.S. Agency for International Development (USAID) by the Quality Assurance Project.

Boyatzis, R.E. 1982. *The Competent Manager: A Model for Effective Performance.* New York. Wiley.

Brown, L.P., L.M. Franco, N. Rafeh, and T. Hatzell. 1992. *Quality assurance of healthcare in developing countries.* Bethesda, MD: Report published for the U.S. Agency for International Development (USAID) by the Quality Assurance Project.

Bryce, J., A. Voigt, A. Adegorovye, B. Zakari, D. Oyebolu, A. Rodman, and S. Saba. n.d. *Skills assessment in primary healthcare training.* (Report for the Africa Regional Project, 698-0421). Washington, DC: United States Agency for International Development.

Butler, F.C. 1978. The concept of competence: An operational definition. *Educational Technology* January, 7–18.

Cabana, M.D., C.S. Rand, N.R. Powe, A.W. Wu, M.H. Wilson, P. Abbond, and H.R. Rubin. 1999. Why don't physicians follow clinical practice guidelines? A framework for improvement. *Journal of the American Medical Association* 282(15):1458–65.

Campbell, J.P., R.A. McCloy, S.H. Oppler, and C.E. Sager. 1993. A theory of performance. In *Personnel Selection in Organizations,* eds. N. Schmitt, W.C. Borman, and Associates. San Francisco: Jossey-Bass.

Carper, B.A. 1992. Fundamental patterns of knowing in nursing. In *Perspectives on Nursing Theory,* ed. L.H. Nicoll. Philadelphia: J.B. Lippincott Company.

Centra Health. 1999. Competency Assessment Model. Lynchburg, Virginia.

CHR (Center for Health Research). 1999. Quality assessment study: Service delivery expansion support project, Indonesia. Jakarta: University of Indonesia.

Coates, V.E., and M. Chambers. 1992. Evaluation of tools to assess clinical competence. *Nurse Education Today* 12:122–29.

Cockerill, R., and J. Barnsley. 1997. Innovation theory and its application to our understanding of the diffusion of new management practices in healthcare organizations. *Healthcare Management Forum* 10:35–8.

Cohen, S.J., H.W. Halvorson, and C.A. Gosselink. 1994. Changing physician behavior to improve disease prevention. *Preventive Medicine* 23:284–91.

Cohen, D.S., J.A. Colliver, M.S. Marcey, E.D. Fried, and M.H. Swartz. 1996. Psychometric properties of a standardized-patient checklist and rating-scale form used to assess interpersonal and communication skills. *Academic Medicine Supplement* 71:S87–S89.

Colliver, J.A., N.V. Vu, and H.S. Barrows. 1992. Screening test length for sequential testing with a standardized-patient examination: A receiver operating characteristic analysis. *Academic Medicine* 67:592–95.

Colliver, J.A., and R.G. Williams. 1993. Technical issues: Test application. *Academic Medicine* 68:454–59.

Connor, M.A., and A.F. Cook. 1995. Laser operator training and competency assessment in the community hospital setting. *Journal of Laser Applications* 7:177–81.

Cusimano, M.D., R. Cohen, W. Tucker, J. Murnaghan, R. Kodama, and R. Reznick. 1994. A comparative analysis of the costs of administration of an OSCE. *Academic Medicine* 69:571–76.

Das, M., D. Mpofu, E. Dunn, J.H. Lanphear. 1998. Self and tutor evaluations in problem-based learning tutorials: Is there a relationship? *Medical Education* 32:411–18.

Davis, D.A., M.A. Thomas, A.D. Oxman, and R.B. Haynes. 1992. Evidence for the effectiveness of CME: A review of 50 randomized controlled trials. *Journal of the American Medical Association* 268(9):1111–17.

———. 1995. Changing physician performance: A systematic review of the effect of continuing medical education strategies. *Journal of the American Medical Association* 274(9):700–04.

Devitt, P., and E. Palmer. 1998. Computers in medical education 3: A possible tool for the assessment of clinical competence? *Australian and New Zealand Journal of Surgery* 68:602–04.

Duller, S.L.F. 1995. Determining competence of experienced critical care nurses. *Nursing Management* 26:48F–48H.

Elnicki, D.M., W.T. Shockcor, D.K. Morris, and K.A. Halbritter. 1993. Creating an objective structured clinical examination for the internal medicine clerkship: Pitfalls and benefits. *The American Journal of the Medical Sciences* 306:94–97.

Ericsson, K.A., ed. 1996. *The Road to Excellence.* Mahwah, NJ: Lawrence Erlbaum Associates.

Erez, M. 1990. Performance quality and work motivation. In *Work Motivation,* eds. U. Kleinbeck, H. Quast, H. Thierry, and H. Hacker. Hillsdale, NJ: Erlbaum.

Fenton, M.V. 1985. Identifying competencies of clinical nurse specialists. *Journal of Nursing Administration* 15:32–37.

Fincher, R.E., K.A. Lewusm, and T.T. Kuske. 1993. Relationships of interns' performance to their self-assessments of their preparedness for internship and to their academic performance in medical school. *Academic Medicine Supplement* 68:S47–S50.

Fitzpatrick, J.M., A.E. While, and J.D. Roberts. 1994. The measurement of nurse performance and its differentiation by course of preparation. *Journal of Advanced Nursing* 20:761–68.

Fleishman, E.A., and C.J. Bartlett. 1969. Human abilities. *Annual Review of Psychology* 20:349–80.

Franco, L.M., C. Franco, N. Kumwenda, and W. Nkloma. 1996. Comparison of methods for assessing quality of health worker performance related to management of ill children. *Quality Assurance Methodology Refinement Series.* Bethesda, MD: Published for the U.S. Agency for International Development (USAID) by the Quality Assurance Project.

Franco, L.M., C.C. Daly, D. Chilongozi, and G. Dallabetta. 1997. Quality of case management of sexually transmitted diseases: Comparison of the methods for assessing the performance of providers. *Bulletin of the World Health Organization* 75(6):523–32.

Franco, L.M., R. Kanfer, L. Milburn, R. Qarrain, and P. Stubblebine. 2000. An in-depth analysis of individual determinants and outcomes of health worker motivation in two Jordanian hospitals. Bethesda, MD: Partnership for Health Reform Project-Abt Associates.

Friedmanc, C.P., A.S. Elstein, F.M. Fredric, G.C. Murphy, T.M. Franz, P.S. Heckerling, P.L. Fine, T.M. Miller, V. Abraham. 1999. Enhancement of clinicians' diagnostic reasoning by computer-based consultation. *Journal of American Medical Association* 282:1851–56.

Gordon, J.J., N.A. Saunders, D. Hennrikus, and R.W. Sanson-Fisher. 1992. Interns' performance with simulated patients at the beginning and the end of the intern year. *Journal of General Internal Medicine* 7:57–62.

Green, L., M. Kreuter, S. Deeds, K. Partridge. 1980. *Health Education Planning: A Diagnosis Approach.* Palo Alto, CA: Mayfield Press.

Hamric, A.B. 1989. History and overview of the CNS role. In *The clinical nurse specialist in theory and practice,* eds. A.B. Hamric and J.A. Spross. Philadelphia: W.B. Saunders Co.

Harden, R.M., and F.A. Gleeson. 1979. Assessment of clinical competence using an objective structured clinical examination (OSCE). *Medical Education* 13:41–54.

Hebers, J.E., G.L. Noel, G.S. Cooper, J. Harvey, L.N. Pangaro, and M.J. Weaver. 1989. How accurate are faculty evaluations of clinical competence? *Journal of General Internal Medicine* 4:202–08.

Hermida, J., D. Nicholas, and S. Blumenfeld. 1996. Comparative Validity of Three Methods for Assessment of Quality of Primary Healthcare: Guatemala Field Study. *Quality Assurance Methodology Refinement Series.* Bethesda, MD: Published for the U.S. Agency for International Development (USAID) by the Quality Assurance Project.

Hodges, B., G. Regehr, N. McNaughton, R. Tiberius, M. Hansom. 1999. OSCE checklists do not capture increasing levels of expertise. *Academic Medicine* 74:1129–34.

Hofstede, G. 1980. *Culture's Consequences: International Differences in Work-Related Values.* Beverly Hill, CA: Sage.

Hull, A.L., S. Hodder, B. Berger, D. Ginsberg, N. Lindheim, J. Quan, and M.E. Kleinhenz. 1995. Validity of three clinical performance assessments of internal medicine clerks. *Academic Medicine* 70:517–22.

Irvine, D. 1997. The performance of doctors. I: Professionalism and self regulation in a changing world. *British Medical Journal* 314:1540–02.

Irvine, D. 1997. The performance of doctors. II: Maintaining good practice, protecting patients from poor performance. *British Medical Journal* 314:1613.

Jaeger, A.M. 1990. The applicability of western management techniques in developing countries: A cultural perspective. In *Management in Developing Countries,* eds. A.M. Jaeger and R.N. Kanungo. London: Routledge.

Jansen, J.J.M., L.H.C. Tan, C.P.M. Van der Vleuten, S.J. Van Luijk, J.J. Rethans, and R.P.T.M. Grol. 1995. Assessment of competence in technical clinical skills of general practitioners. *Medical Education* 29:247–53.

JCAHO (Joint Commission on Accreditation of Healthcare Organizations). 1996. Comprehensive accreditation manual for hospitals: The official handbook. Oakbrook, IL.

Kanungo, R.N. and A.M. Jaeger. 1990. Introduction: The need for indigenous management in developing countries. In *Management in Developing Countries,* in A.M. Jaeger and R.N. Kanungo. London: Routledge.

Kapil, U., A.K. Sood, D. Nayar, D.R. Gaur, D. Paul, S. Chaturvedi, and M. Srivasta. 1992. Assessment of knowledge and skills about growth monitoring amounts medical officers and multi purpose workers. *Indian Pediatrics* 31:43–46.

Kelley, E., C. Geslin, S. Djibrina, and M. Boucar. 2000. The impact of QA methods on compliance with the Integrated Management of Childhood Illness algorithm in Niger. *Operations Research Results* 1(2). Bethesda, MD: Published for the U.S. Agency for International Development (USAID) by the Quality Assurance Project.

Kelly, M.H., L.M. Campbell, and T.S. Murray. 1999. Clinical skills assessment. *British Journal of General Practice* 49:447–50.

Kim, Y.M., F. Putjuk, A. Kols, and E. Basuki. 2000. Improving provider-client communication: Reinforcing IPC/C training in Indonesia with self-assessment and peer review. *Operations Research Results* 1(6). Bethesda, MD: Published for the U.S. Agency for International Development (USAID) by the Quality Assurance Project.

Kinnersley, P., and R. Pill. 1993. Potential of using simulated patients to study the performance of general practitioners. *British Journal of General Practice* 43:297–300.

Klein, R. 1998. Competence, professional self regulation and the public interest. *British Medical Journal* 316:1740–42.

Klimoski, R.J., and M. London. 1974. Role of the rater in performance appraisal. *Journal of Applied Psychology* 59:445–51.

Knebel. E. 2000. The use and effect of computer-based training: What do we know? *Operations Research Issue Papers* 1(2) Bethesda, MD: Published for the U.S. Agency for International Development (USAID) by the Quality Assurance Project.

Kramer, D. 1996. Keeping competency up when patient numbers are down. *Kansas Nurse* 71:4–5.

Landy, F.J. 1985. *Psychology of Work Behavior.* (3rd ed.). Homewood, IL: Dorsey Press.

Lane, D.S., and V.S. Ross. 1998. Defining competencies and performance indicators for physicians in medical management. *American Journal of Preventive Medicine* 14:229–36.

Latham, G.P. 1986. Job performance and appraisal. In *International Review of Industrial and Organizational Pychology,* eds. C.L. Cooper and I. Robertson. New York: Wiley.

Lenburg, C.B. 1999. Redesigning expectations for initial and continuing medical competence for contemporary nursing practice. *Online Journal of Issues in Nursing.*

Locke, E.A., and G.P. Latham. 1990. Work motivation: The high performance cycle. In *Work Motivation,* eds. U. Kleinbeck, H. Quast, H. Thierry, and H. Hacker. Hillsdale, NJ: Erlbaum.

Loevinsohn, B.O., E.T. Gierrero, and S.P. Gregorio. 1995. Improving primary healthcare through systematic supervision: A controlled field trial. *Health Policy and Planning* 10:144–53.

Long, H.B. 1989. Selected principles developing self-direction in learning. Paper presented at the Annual Meeting of the American Association for Adults and Continuing Education. Atlantic City, NJ (October).

Lyons, T.F. 1974. The relationship of physicians' medical recording performance to their medical care performance. *Medical Care* 12:463–69.

Main, D.S., S.J. Cohen, and C.C. DiClemente. 1995. Measuring physician readiness to change cancer screening: Preliminary results. *American Journal of Preventive Medicine* 11:54–58.

MacDonald, P. 1995. *The peer review program of the Indonesian midwives association* (Final Report of Phase Two of the Pilot Project). Bethesda, MD: University Research Corporation.

Marquez, Lani. Forthcoming. Helping Healthcare Providers Perform According to Standards. *Operations Research Issue Paper.* Bethesda, MD: To be published for the U.S. Agency for International Development (USAID) by the Quality Assurance Project.

McCaskey, L., and M. LaRocco. 1995. Competency testing in clinical microbiology. *Laboratory Medicine* 26:343–49.

McFaul, P.B., and P.W. Howie. 1993. The assessment of clinical competence in obstetrics and gynecology in two medical schools by an objective structured clinical examination. *British Journal of Obstetrics and Gynecology* 100:842–46.

Mendonca, M., and R.N. Kanungo. 1994. Motivation through effective reward management in developing countries. In *Work Motivation. Models for Developing Countries,* eds. R.N. Kanungo and M. Mendonca. New Delhi: Sage Publications.

Miller, G.E. 1993. Conference summary. *Academic Medicine* 68:471–76.

Misra, S., and R.N. Kanungo. 1994. Bases of work motivation in developing societies: A framework for performance management. I n *Work Motivation. Models for Developing Countries,* eds. R.N. Kanungo and M. Mendonca. New Delhi: Sage Publications.

Mittman, B.S., X. Tonesk, and A. Jacobson. 1995. Implementing clinical practice guidelines: Social influence strategies and practitioner behavior change. *Quality Review Bulletin* 18:413–22.

Mohrman, A., S.M. Resnick-West, and E.E. Lawler. 1989. *Designing performance appraisal systems.* San Francisco: Jossey-Bass.

Mumford, M.D. 1983. Social comparison theory and the evaluation of peer evaluations: A review and some applied implications. *Personnel Psychology* 36:867–81.

Newble, D.I., and D.B. Swanson. 1988. Psychometric characteristics of the objective structured clinical examination. *Medical Education* 22:335–41.

Nicholls, J.G. 1984. Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychological Review* 91:328–46.

Nicholls, J.G., and A.T. Miller. 1984. Development and its discontents: The differentiation of the concept of ability. In *Advances in motivation and achievement. The development of achievement motivation* (Vol. 3), ed. J.G. Nicholls, Greenwich, CN: JAI Press.

Norman, G.R., V.R. Neufeld, A. Walsh, C.A. Woodward, and G.A. McConvey. 1985. Measuring physicians' performances using simulated patients. *Journal of Medical Education* 60:925–34.

Norman, G.R., C.P.M. Van der Vleuten, and E. DeGraaff. 1991. Pitfalls in the pursuit of objectivity: Issues of validity, efficiency and acceptability. *Medical Education* 25:119–26.

Ofori-Adjei, D., and D.K. Arhuinful. 1996. Effect of training on the clinical management of malaria by medical assistants in Ghana. *Social Science Medicine* 42:1169–76.

Peabody, J.W., J. Luck, P. Glassman, T.R. Dresselhaus, M. Lee. 2000. Comparison of vignettes, standardized patients, and chart abstraction: A prospective validation study of 3 methods for measuring quality. *Journal of American Medical Association* 283:1715–22.

Pietroni, M. 1995. The assessment of competence in surgical trainees. *Surgical Training* 200–02.

Pritchard, R.D. 1990. Enhancing work motivation through productivity measurement and feedback. In *Work Motivation,* eds. U. Kleinbeck, H. Quast, H. Thierry, and H. Hacker. Hillsdale, NJ: Erlbaum.

Prochaska, J.O., and C.C. DiClemente. Toward a comprehensive model of change. In *Treating Addictive Behaviors,* eds. W.R. Miller and N. Heather. New York: Plenum, 1986.

Ram, P., C. Van der Vleuten, J.J. Rethans, J. Grol, and K. Aretz. 1999. Assessment of practicing family physicians: Comparison of observation in a multiple-station examination using standardized patients with observation of consultations in daily practice. *Academic Medicine* 74:62–69.

Ready, R.W. 1994. Clinical competency testing for emergency nurses. *Journal of Emergency Nursing* 20:24–32.

Reznick, R., S. Smee, A. Rothman, A. Chalmers, D. Swanson, L. Dufresne, G. Lacombe, J. Baumber, P. Poldre, L. Levasseur, R. Cohen, J. Mendez, P. Patey, D. Boudreau, and M. Berard. 1992. An objective structured clinical examination for the licentiate: Report of pilot project of the Medical Council of Canada. *Academic Medicine* 67:487–94.

Rogers, E.M. 1962. *Diffusion of Innovations.* New York, NY: The Free Press.

Rohwer, J., and B. Wandberg. 1993. *An Innovative School Health Education Model Designed for Student Achievement.*

Salazar-Lindo, E., E. Chea-Woo, J. Kohatsu, and P.R. Miranda. 1991. Evaluation of clinical management training programme for diarrhoea. *Journal of Diarrheoal Disease Research* 9:227–34.

Sandvik, H. 1995. Criterion validity of responses to patient vignettes: An analysis based on management of female urinary incontinence. *Family Medicine* 6:388–92.

Schuler, S.R., E.N. McIntosh, M.C. Goldstein, and B.R. Pandi. 1985. Barriers to effective family planning in Nepal. *Studies in Family Planning* 16:260–70.

Sloan, D.A., M.B. Donnelly, S.B. Johnson, R.W. Schwartz, and W.E. Strodel. 1993. Use of an objective structured clinical examination (OSCE) to measure improvement in clinical competence during the surgical internship. *Surgery* 114:343–51.

Smith, J.E., and S. Merchant. 1990. Using competency exams for evaluating training. *Training and Development Journal* 65–71.

Snyder, W., and S. Smit. 1998. Evaluating the evaluators: Interrater reliability on EMT licensing examinations. *Prehospital Emergency Care* 2:37–47.

Southgate, L., and D. Dauphinee. 1998. Maintaining standards in British and Canadian medicine: The developing role of the regulatory body. *British Medical Journal* 316:697–700.

Spencer, L.M., D.C. McClelland, and S.M. Spencer. 1994. *Competency assessment methods.* Boston: Hay/McBer Research Press.

Spross, J.A., and J. Baggerly. 1989. Models of advanced nursing practice. In *The Clinical Nurse Specialist in Theory and Practice,* eds. A.B. Hamric and J.A. Spross. Philadelphia: W.B. Saunders Company.

Steiner, B.D., R.L. Cook, A.C. Smith, P. Curtis. 1998. Does training location influence the clinical skills of medical students? *Academic Medicine* 73:423–26.

Stillman, P.L., D. Swanson, S. Smee, A.E. Stillman, T.H. Ebert, V.S. Emmel, J. Caslowitz, H.L. Greene, M. Hamolsky, C. Hatem, D.J. Levenson, R. Levin, G. Levinson, B. Ley, G.J. Morgan, T. Parrino, S. Robinson, and J. Willms. 1986. Assessing clinical skills of residents with standardized patients. *Annals of Internal Medicine* 105:762–71.

Stillman, P, D. Swanson, M.B. Regan, M.M. Philbin, V. Nelson, T. Ebert, B. Ley, T. Parrino, J. Shorey, A. Stillman, E. Alpert, J. Caslowitz., D. Clive, J. Florek, M. Hamolsky, C. Hatem, J. Kizirian, R. Kopelman, D. Levenson, G. Levinson, J. McCue, H. Pohl, F. Schiffman, J. Schwartz, M. Thane, and M. Wolf. 1991. Assessment of clinical skills of residents utilizing standardized patients. *Annals of Internal Medicine,* 114:393–401.

Stillman, P.L. 1993. Technical issues: Logistics. *Academic Medicine* 68:464–68.

Stoy, W.A. n.d. EMT-paramedic and EMT-intermediate continuing education: National guidelines. National Highway Traffic Safety Administration.

Sullivan, C.A. 1994. Competency assessment and performance improvement for healthcare providers *Journal of Health Quarterly* 16:14–19.

Sullivan, R.L. 1996. Transferring performance skills: A clinician's case study. *Technical & Skills Training* 14–16.

Swartz, M.H., J.A. Colliver, C.L. Bardes, R. Charon, E.D. Fried, and S. Moroff. 1999. Global ratings of videotaped performance versus global ratings of actions recorded on checklists: A criterion for performance assessment with standard patients. *Academic Medicine* 74(9):1028–32.

Tamblyn, R.M., D.J. Klass, G.K. Schnabl, and M.L. Kopelow. 1991. The accuracy of standardized patient presentation. *Medical Education* 25:100–09.

Upmeyer, A. 1971. Social perception and signal detectability theory: Group influence on discrimination and usage of scale. *Psychol Forsch* 34:283–94.

Van der Vleuten, C.P.M., and D.B. Swanson. 1990. Assessment of clinical skills with standardized patients: State of the art. *Teaching and Learning in Medicine* 2:58–76.

Van der Vleuten, C.P.M., G.R. Norman, and E. De Graaff. 1991. Pitfalls in the pursuit of objectivity: Issues of reliability. *Medical Education* 25:110–18.

Velicer, W.F., J.S. Rossi, C.C. DiClemente, and J.O. Prochaska. 1996. A criterion measurement model for health behavior change. *Addictive Behavior* 5:555–84.

Vu, N.V., H.S. Barrows, M.L. March, S.J. Verhulst, J.A. Colliver, and T. Travis. 1992. Six years of comprehensive, clinical, performance-based assessment using standardized patients at the Southern Illinois University School of Medicine. *Academic Medicine* 67:42–49.

While, A.E. 1994. Competence versus performance: Which is more important? *Journal of Advanced Nursing* 29:525–31.

Wood, E.P., and E. O'Donnell. 2000. Assessment of competence and performance at interview. *British Journal of Medicine* 320:S2–7231.

Wood, R., and C. Power. 1987. Aspects of the competence-performance distinction: Educational, psychological and measurement issues. *Journal of Curriculum Studies* 19:409–24.